Development of Framework for Identifying Mobility Desert

or

Identifying multi-modal deserts: a multivariate outlier detection approach

Center for Transportation, Environment, and Community Health
Final Report

by
Yu Zhang, Peng Chen, Yujie Guo

February 16, 2021

DISCLAIMER

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2.Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Identifying multi-modal deserts: a multivariate outlier detection approach		5. Report Date February 16, 2021
		6. Performing Organization Code
7. Author(s) Yu Zhang (ORCID: 0000-0003-1202-626X) Peng Chen (ORCID: 0000-0002-3201-7339) Yujie Guo (ORCID: 0000-0003-3820-6840)		8. Performing Organization Report No.
9. Performing Organization Name and Address Department of Civil and Environmental Engineering Department of Urban and Regional Planning University of South Florida, FL 33620		10. Work Unit No.
		11. Contract or Grant No. 69A3551747119
12. Sponsoring Agency Name and Address U.S. Department of Transportation 1200 New Jersey Avenue, SE Washington, DC 20590		13. Type of Report and Period Covered Final Report 10/01/2020 – 12/31/2021
		14. Sponsoring Agency Code US-DOT
15. Supplementary Notes		

16. Abstract

Providing all means of travel facilitates people's access to jobs, healthcare, critical activities, and other services. To enable equal multi-modal mobility services to the public, it is important to evaluate equity in accessing different travel modes. In this study, we proposed a concept called "multi-modal deserts" and developed an approach to identify them. Multi-modal deserts refer to areas with limited mobility services that constrain people from accessing services and opportunities. Framed under multi-modality, multivariate outlier detection was applied to identify areas' mobility services that significantly deviate from other areas by analyzing road network factors and travel modes. Downtown Tampa, Florida, was selected as an empirical case to demonstrate the proposed method, and 11 multi-modal deserts were identified among 182 Census Block Groups. In addition, spider charts were used to illustrate and compare the features of these multi-modal deserts. The results show that two multi-modal deserts in central Downtown Tampa have the highest poverty ratios and have very limited access to all travel modes. For such multi-modal deserts, transit and shared micromobility need to be better served in a way to enrich the travel mode choices for low-income residents. Other multi-modal deserts are at the edge of Downtown Tampa, which has no access to shared micromobility and limited access to transit. The results will help local authorities identify mobility gaps by better allocating resources and improving equal access to opportunities for all citizens.

17. Key Words		18. Distribution Statement	
Multi-modal desert, Outlier detection, Transportation equity, Mode access		Public Access	
19. Security Classif (of this report)	20. Security Classif. (of this page)	21. No of Pages	22. Price
Unclassified	Unclassified		

Form DOT F 1700.7 (8-69)

ABSTRACT

Providing all means of travel facilitates people's access to jobs, healthcare, critical activities, and other services. To enable equal multi-modal mobility services to the public, it is important to evaluate equity in accessing different travel modes. In this study, we proposed a concept called "multi-modal deserts" and developed an approach to identify them. Multi-modal deserts refer to areas with limited mobility services that constrain people from accessing services and opportunities. Framed under multi-modality, multivariate outlier detection was applied to identify areas' mobility services that significantly deviate from other areas by analyzing road network factors and travel modes. Downtown Tampa, Florida, was selected as an empirical case to demonstrate the proposed method, and 11 multi-modal deserts were identified among 182 Census Block Groups. In addition, spider charts were used to illustrate and compare the features of these multi-modal deserts. The results show that two multi-modal deserts in central Downtown Tampa have the highest poverty ratios and have very limited access to all travel modes. For such multi-modal deserts, transit and shared micromobility need to be better served in a way to enrich the travel mode choices for low-income residents. Other multi-modal deserts are at the edge of Downtown Tampa, which has no access to shared micromobility and limited access to transit. The results will help local authorities identify mobility gaps by better allocating resources and improving equal access to opportunities for all citizens.

INTRODUCTION

Transportation resources are unequally distributed in space, and it has been a convention that transport policy and planning practices are centered on automobiles while ignoring other road users. These underserved road users generally are low-income minorities, and their travel choices rely heavily on transit (*1*). In recent years, there has been a trend of adopting emerging travel modes such as e-scooter- sharing (*2*) and ride-sourcing (*3*). However, when distributing bikes and e-scooters on streets, the location choice of these emerging programs involves selection bias, which raises public debate concerning equity in accessing these shared mobility services. Yet, their prices remain unaffordable, and the locations of e-bike and e-scooter stations remain inaccessible for low-income populations (*4; 5*). People living in underserved communities continually lose opportunities to access destinations and encounter certain levels of socioeconomical segregation.

This study proposed a concept called "multi-modal deserts," referring to areas with limited mobility options and services. Numerous studies have quantified access to a single travel mode (*6*), and related topics include but are not limited to access to transit (*7*), shared bikes (*4; 8*), etc. However, few studies have incorporated multiple travel modes simultaneously. As a transportation system is intrinsically multimodal, evaluating the quality of mobility services in an area and all travel modes should be jointly considered.

Emerging modes such as paratransit provide free services and mitigate mobility challenges faced by vulnerable groups. For example, the Sunshine Line in Hillsborough County (Florida) provides free door-to-door mobility services for older adult and low-income residents and persons with disabilities who cannot physically or economically afford a car. The service helps these vulnerable groups access destinations for critical activities such as healthcare, grocery shopping, and jobs. Shared micromobility modes also help address gaps in the transportation system. In Portland (Oregon) (*9*) and Baltimore (Maryland) (*10*), shared e-scooter companies are required to allocate at least 20 percent (Portland) or 25 percent (Baltimore) of their total fleets to disadvantaged areas. Cities are paying attention to areas with constrained mobility services, which echoes the idea of multi-modal deserts. However, to date, there is no rigorous method to identify multi-modal deserts. Practitioners need methods and tools to identify underserved communities to efficiently allocate resources. Thus, a framework needs to be established to help practitioners and decision-makers identify areas that urgently need public assistance in improving mobility services.

Improving mobility service is about facilitating access to different travel modes and improving the quality of related services (*11*). It is important to consider mode access, as it is a prerequisite for accessibility to key destinations. Measuring the access of different travel modes is related to the supply of infrastructure (*12*). Multi-modal deserts can be interpreted as areas having limited supply for one or multiple type(s) of infrastructure, i.e., the supply of services in multi-modal deserts tends to deviate from the majority of areas, making them outliers. Therefore, outlier detection methods can be applied to identify multi-modal deserts.

This study proposed a framework to detect multi-modal deserts. Procedures include travel mode selection, data preparation, outlier detection, and results interpretation. GIS tools were used for visualization. Downtown Tampa was selected for empirical analysis. Noted as a Sunbelt city, people in Tampa have a high level of car dependency, and transit and other sustainable alternatives are unequally distributed in space. These features make Tampa a representative case for most U.S. cities that are auto-oriented and shifting towards multi-modal transportation.

The contributions of this study are fourfold. First, it defines a new term—multi-modal desert—a concept that is important for addressing the emerging needs of multi-modality and brings benefits to underserved populations. Second, previous studies on transportation equity are unimodal-based; a multi-modal desert considers the quantity and quality of all travel modes regarding their supply, which extends inequality of access to different travel modes to a larger context. Third, this study conceptualizes a framework to identify multi-modal deserts; this approach can be formulated as a tool for identifying underserved areas. Fourth, the empirical results serve to inform policies and practices to address mobility challenges in the local environment and enhance transportation equity.

LITERATURE REVIEW

As noted, the concept of a multi-modal desert is not new; it is an extended form of a transit desert. It evaluates the gap in assessing different travel modes and promotes transportation equity, and it requires new methods for identification. This section discusses access to different travel modes and methods for determining multivariate outliers.

Access to Different Travel Modes

The concept of a multi-modal desert is an extended form of a transit desert, which refers to an area that has inadequate transit services to the public (*13-15*). A multi-modal desert extends a single mode from transit to all travel modes.

To identify a multi-modal desert, it is important to quantify access to different travel modes. According to Van Wee et al. (*12*), transportation accessibility has various definitions that can be grouped into infrastructure-related or activity-oriented approaches. The activity-oriented approach accounts for trip activities (purposes), e.g., if people can reach a destination within a certain time threshold. Infrastructure-related approach is adopted in this study, it is supply-oriented. Its accessibility is measured by characteristics of infrastructure supplies such as length of road network. Considering the supply of transportation modes, the existing literature discusses access to different travel modes independently, such as access to private vehicles, public transit, and shared modes such as Transportation Network Companies (TNCs), bike sharing, and e-scooter sharing.

To measure private vehicle access, the share of the population (or households) that owns a car or has a driver's license can be used as an indicator (*16*). Owning a private car helps economically-disadvantaged populations to access jobs (*17; 18*). Access to transit can be a composite measurement, which is jointly quantified with various transit-related features such as capacity of a transit line, headway, hours of operation, speed, distance to a bus stop, etc.(*19*). This measurement can be computed for each bus stop and can be aggregated to larger geographic scales (e.g., census zones). This measurement helps evaluate transit services across different zones. With regard to shared mobility access, one approach is called coverage-based. Ursaki and Aultman-Hall (*4*) defined coverage/service areas of a bike-sharing station by creating buffers around stations; those who reside within the coverage area are assumed to have access to shared bikes. To inform the practice for transportation equity improvement, socioeconomic factors within and outside the coverage area have been analyzed (*4; 8*). Another approach uses mean density or the availability of shared bikes within a given area to measure the supply and quality of services; areas with higher population density and higher income are related to a higher supply of shared bikes (*20-22*). To measure access to taxi and ride-sourcing services, fleets across spatial spaces can be used as an indicator (*5*). However, ride-sourcing companies (TNCs) such as Uber and Lyft rarely share their operational data with researchers. Equity discussions on TNCs are still limited.

When measuring the service quality of a transportation mode, the quality of road networks and infrastructure also needs to be considered, as they potentially influence people's mode choice for a trip. For example, the availability of sidewalks and bike lanes helps promote active transportation (*23*), and longer bike lanes positively influence ridership of shared bikes (*24*). A study in Portland found that a disproportionate share of bicycling occurs on streets with bike lanes (*25*). In addition, sidewalks and bike lanes are positively related to the use of transit (*26*). Road intersections and road length are important measurements of network connectivity (*26*), which also influence a mode's service quality. A well-connected road network has many short links and intersections that greatly improve accessibility (*27*).

Outlier detection

As noted, the quantity of mobility services in a given area is characterized by transportation-related features. Multi-modal deserts have limited access to transportation modes and supplies, deviating from other areas. Thus, outlier detection methods can be applied to locate multi-modal deserts. This section provides a brief overview of popular outlier detection methods.

Outliers are generally defined as data points that are far outside the norm of variables or populations (*28-30*). They can arise from different mechanisms such as errors in the data (e.g., data misreporting) and inherent variability of the data (*31*). Retaining outliers in data may lead to false statistical inferences (*32*).

Outlier detection is a hot research topic in statistics and related application fields and can be applied to one or multiple variables. Classic approaches for detecting univariate outliers are box-plot (*33*), the Ven der Loo method (*34*), etc. For multi-dimensional data, many advanced outlier detection methods have been developed and can be categorized as statistical models, neural networks, or machine learning algorithms (*35*). A comprehensive exploration of outlier detection methods can be found in the existing literature (*35-37*). Given the tremendous choices of outlier detection methods, Hodge and Austin noted that the selected methodology should accurately model the data distribution and attribute types of data, scalability, and speed (*35*). Among outlier detection methods, the most classical approach for multivariate outlier detection is Mahalanobis distance (*38; 39*), (*40*), which measures the number of standard deviations that the value of an observation deviates from the mean considering the multivariate covariance structure. Applying a Mahalanobis distance assumes that observations follow a multivariate normal distribution; thus, the distribution of Mahalanobis distance follows a chi-square distribution (*41*). An outlier is identified by comparing the Mahalanobis distance of an observation with a critical value of the χ^2 distribution with p degrees of freedom; a 97.5 percent significance level is often suggested (*42*). Applying Mahalanobis distance requires the estimation of mean and covariance. Classic estimations use all sample data that can be influenced by outlying observations (*43*). Thus, classic estimation based on Mahalanobis distance can be unreliable if outliers are presented in the data. To get an unbiased sample mean and covariance, robust estimations are needed. A popular robust estimation is using a minimum covariance determinant (MCD) estimator (*44; 45*), which aims to find a subset of observations that have the smallest determinant of a sample covariance matrix. Mahalanobis distance also suffers from the "curse of dimensionality" caused by high dimensional features, i.e., data are spread through a larger volume and become less dense, making it more difficult to discern the outliers in the full space rather than its subspace (*35; 46; 47*). To address this issue, principal component analysis (PCA) can be used for dimension reduction (*35*). PCA creates a linear combination of original values, referred to as a dataset, reorganized by a series of principal components (PCs). PCs are orthogonal to each other and retain the most variability in the data (*48*). The distance measures (e.g., Mahalanobis distance, score distance, orthogonal distance) (*49*) can be applied to the first k PCs to unmask outlying observations. Similar to the Mahalanobis distance approach, computing PCs also requires the estimation of a covariance matrix, and a robust estimation (e.g., MCD) must be applied. Prior research compared the results of different robust PCA-based approaches using simulated data, and ROBPCA (*50*) performed the best (*49*). This study adopted the ROBPCA method to detect potential outliers.

METHODOLOGY

This section includes an introduction to data collection and a detailed description of the method.

Data collection

Table 1 describes how selected variables were measured or calculated, with data sources listed. As shown, most targeted variables were included except TNCs due to difficulty in accessing their data. Public transit connectivity is a composite measurements, and calculation of the index refers to the work of Mishra, Welch, and Jha (*19*).

TABLE 1: Description of Variables

Variable	Description	Source
Road network		
Road centerline length	Total length of road centerline length inside a Census Block Group scaled by area size	City of Tampa GeoHub (2020)
Number of intersections	Number of intersections inside a Census Block Group scaled by total length of road centerline	City of Tampa GeoHub (2020)
Sidewalk length	Total length of sidewalks inside a Census Block scaled by total length of road centerline	City of Tampa GeoHub (2020)
Bike lane length	Total length of bike lanes inside a Census Block divided by total length of road centerline	City of Tampa GeoHub (2020)
Travel mode-related features		
Car ownership percent	Percentage of households own private vehicles in a Census Block Group	TBRPM (2015)
Transit connectivity	Connectivity of transit nodes in a Census Block Group	HART (2020)
Shared bike counts	Shared bike counts in a Census Block Group scaled by population count in the Census Block Group	City of Tampa (2019)
Shared e-scooter counts	Shared e-scooters counts in a Census Block Group scaled by population count in the Census Block Group	Populus (2019)

Note: Methods to calculate transit connectivity refer to (*19*)

Methodology

The methodological framework adopts four steps to identify multi-modal deserts; the process is presented in Figure 1.

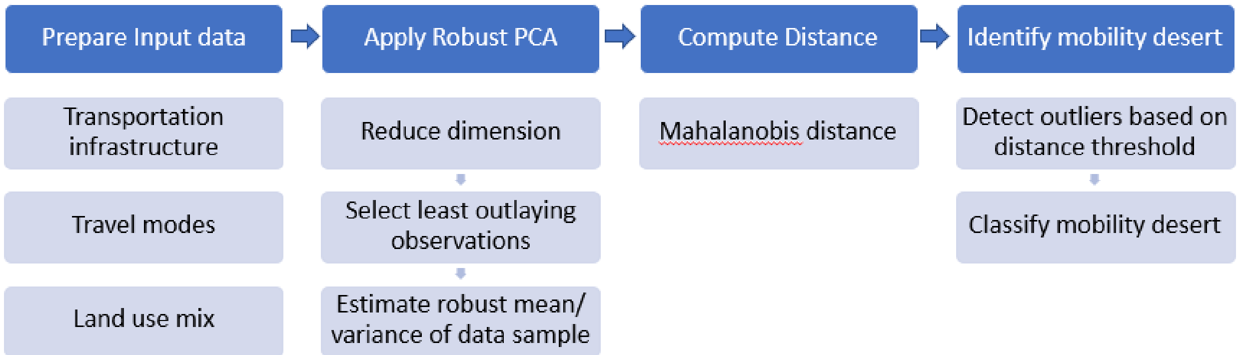


FIGURE 1: Analytical process

Details of this analytical process are described as follows:

Step 1: Prepare input data – This step defines mobility data as a matrix, $\boldsymbol{X} = (X_{i1}, X_{i2}, \ldots X_{ip})$ for $i = 1, \ldots n$. The data have n Census Block Groups with p transportation feature dimensions, with expected value $\boldsymbol{\mu} = (\mu_1, \mu_2 \ldots \mu_p)$.

Step 2: Apply Robust PCA – Computing principal components requires an estimation of covariance matrix, and a classic PCA uses covariance of a sample that can be influenced by outliers. To overcome this limitation, a robust PCA method called ROBPCA (*50*) was adopted in this study. The ROBPCA algorithm consists of three major steps:

1) Given $\boldsymbol{X}_{n,p}$ is a matrix with n Census Block Groups and p transportation features. The first step restricts the feature space up to n by the singular value decomposition (SVD) method, the SVD works follows:

$$\boldsymbol{X}_{n,p} - \boldsymbol{1}_n \widehat{\boldsymbol{\mu}}_0' = \boldsymbol{U}_{n,r_0} \boldsymbol{D}_{r_0,r_0} \boldsymbol{V}^t_{r_0,p}{}' \tag{1}$$

Where $\widehat{\boldsymbol{\mu}}_0'$ is the classic mean vector, $r_0 = rank(\boldsymbol{X}_{n,p} - \boldsymbol{1}_n \widehat{\boldsymbol{\mu}}_0')$. Without the loss of information, the dataset becomes

$$\boldsymbol{X}_{n,r_0} = \boldsymbol{UD} \tag{2}$$

This step is a dimension reduction tool when p ≥ n and also helps remove redundant dimensions when the rank of features is less than p.

2) The second step obtains the preliminary subspace of dimension k_0 for $\boldsymbol{X}_{n,r_0}$. To estimate the subspace, this step first finds h "least outlying" observations by measuring the outlyingness of each observation. The computation applies Stahel-Donoho affine-invariant outlyingness (*51; 52*). The least h outlying observations are selected, and their means and variances are denoted as $\widehat{\boldsymbol{\mu}_1}$ and $\boldsymbol{S}_0$. The covariance matrix $\boldsymbol{S}_0$ is used to decide the number of PCs $k_0 < r_0$ that can be retained in the further analysis. In detail, the spectral decomposition of $\boldsymbol{S}_0$ is conducted as

$$\boldsymbol{S}_0 = \boldsymbol{P}_0 \boldsymbol{L}_0 \boldsymbol{P}_0{}' \tag{3}$$

where $\boldsymbol{L}$ is the eigenvalue and $\boldsymbol{P}$ is the eigenvectors, the eigenvalues $\widetilde{l_J}$ are sorted in descending order and eigenvector is indexed accordingly. The first k_0 PCs ($\boldsymbol{P}_{r_1,k_0}$) are selected. k_0 is chosen by a criterion:

$$\frac{\sum_{j=1}^{k_0} \widetilde{l_J}}{\sum_{j=1}^{r} \widetilde{l_J}} \geq 90\%$$

Projecting $\boldsymbol{X}_{n,r_0}$ to the subspace spanned by the first k_0 of $\boldsymbol{S}_0$ is computed as

$$\boldsymbol{X}_{n,k_0}{}^{*} = (\boldsymbol{X}_{n,r_0} - 1_n \widehat{\boldsymbol{\mu}}_1')\boldsymbol{P}_{r_1,k_0} \tag{4}$$

3) This step robustly estimates the mean and variance of scatter matrix of $\boldsymbol{X}_{n,k_0}{}^{*}$. This step first adopts the FAST-MCD algorithm proposed by Rousseeuw and Van Diressen (1999) to obtain the robustly estimated $\widehat{\boldsymbol{\mu}}_2'$ and $\boldsymbol{S}_1$. Based on $\widehat{\boldsymbol{\mu}}_2'$ and $\boldsymbol{S}_1$, mean $\widehat{\boldsymbol{\mu}}_3'$ and covariance $\boldsymbol{S}_2$ are computed by applying the re-weighted MCD estimator to improve statistical efficiency. The spectral decomposition of $\boldsymbol{S}_1$ can be written as

$$\boldsymbol{S}_1 = \boldsymbol{P}_1 \boldsymbol{L}_1 \boldsymbol{P}_1{}' \tag{5}$$

Where $\boldsymbol{L}_1$ is the diagonal matrix with eigenvalues, and $\boldsymbol{P}_1$ contains the corresponding eigenvectors. The final principal scores are

$$T_{n,k} = (\boldsymbol{X}_{n,k} - \mathbf{1}_n \widehat{\boldsymbol{\mu}}_3')\boldsymbol{P}_1 \tag{6}$$

Step 3: Compute distance – The popular distance measure for detecting outliers is Mahalanobis distance. Mathematically, Mahalanobis distance is defined as

$$MD_{\mu,\Sigma}(\boldsymbol{X}_i) = \sqrt{(\boldsymbol{T}_{n,k} - \boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{T}_{n,k} - \boldsymbol{\mu})} \tag{7}$$

Where $\boldsymbol{T}_{n,k}$ is the vector of PCs obtained from Step 2, and mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ are estimated from PCs. The squared Mahalanobis distance follows a chi-square distribution with *k* degrees of freedom.

Step 4: Identify the multi-modal desert – Given the Mahalanobis distance of an observation (Census Block Group), the outlying area can be identified if

$$MD(\boldsymbol{X}_i) > \sqrt{\chi^2_{k;0.95}} \tag{8}$$

Here, a 95% significance level is selected to account for both multi-modal deserts and mobility advantaged areas. Based on the selected outliers, the standardized feature values are summed for each outlying area. If the sum is lower than 0, the area is considered to be a potential multi-modal desert.

STUDY AREA, RESULTS, AND DISCUSSION

This section describes the study area, presents statistics of features related to mobility services, and discusses multi-modal deserts obtained from the analysis.

Study area: Downtown Tampa

Tampa is located in west central Florida (Figure 2) and has a metropolitan (Tampa-St. Peterburg-Clearwater) population that ranks No. 2 in the state; it is one of the most population-dense metro areas in the US. Tampa provides a variety of mobility options for travelers, including transit, bike sharing, e-scooter sharing, streetcar, water taxi, etc. However, as a Sunbelt city, people living in Tampa have a high level of car dependency. According to DATA USA (2018), for commuting trips, modal splits in Tampa are drive-alone (76.1%), carpool (9.97%), and transit (2.47%). This study selected Downtown Tampa for empirical analysis, as it is a representative case for most US cities labeled with high car dependency and is a growing metro area with new travel options emerging and on track for developing a multi-modal transport system.

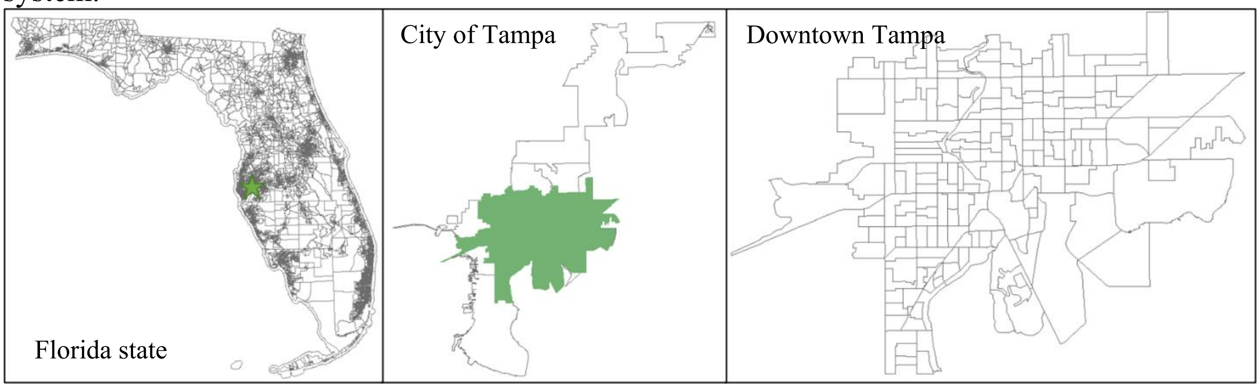


FIGURE 2: Study area

Variable statistics

To conduct the analysis, various data were collected. To better match the US census, the Census Block Group level was selected as the analytical unit. Matching the census can facilitate statistical analysis of disparities in mobility services among different socioeconomic groups. In total, 182 Census Block Groups were included in the study.

Table 2 presents descriptive statistics of the study area, and Figure 3 shows the distribution of transportation resources in Downtown Tampa. Areas with red boundaries are the selected multi-modal deserts discussed later. As shown, most transportation resources are distributed in the central Downtown area, especially bike-sharing and e-scooter-sharing. Features such as road centerline, bike lane, and transit resources also present great variation and have higher value in the Downtown core, leaving some areas underserved. Car ownership presents the opposite variation and is more advantaged around the Downtown edge. People living in central Downtown have better transit and shared micromobility services and, thus, are less dependent on private cars.

TABLE 2: Variable Statistics

Variable	Minimum	Median	Mean	Maximum
Road network				
Road centerline length	92.86	344.90	344.38	968.50
Number of intersections	1.57	6.18	6.33	12.30
Sidewalk length	0.02	1.37	1.30	2.96
Bike lane length	0.00	0.33	0.52	4.20
Transportation mode-related features				
Car ownership percent	0.44	0.88	0.84	1.00
Transit connectivity	0.00	105.19	131.85	779.02
Shared bike counts	0.00	0.00	0.00	0.07
Shared e-scooter counts	0.00	0.00	0.01	0.20

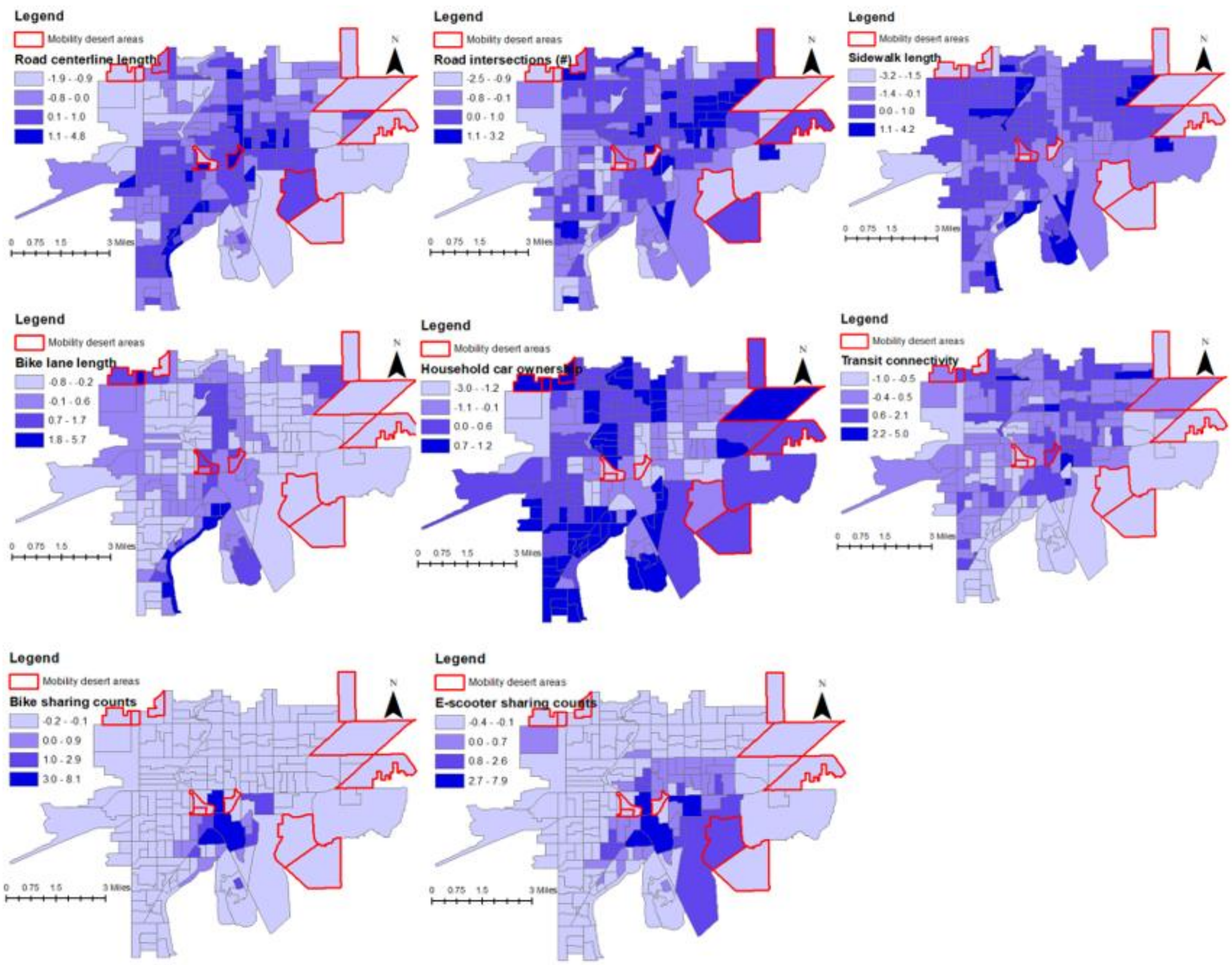


FIGURE 3: Spatial distribution of transportation resources

Results

By following the methodological framework, 11 multi-modal deserts were identified and are presented in Figure 4. As shown, multi-modal deserts are located at both the Downtown edge and neighborhoods near the center of Downtown. As transportation services are centered in the central Downtown area, people living in the edge areas are generally more reliant on private vehicles. Despite central Downtown being equipped with more infrastructure, poor neighborhoods near the central Downtown area, which are just a couple of blocks away, face difficulties in accessing sustainable alternatives, resulting in multi-modal deserts near the central Downtown areas.

The colored areas in Figure 4 indicate the variation of poverty levels in different Block Groups, based on data obtained from the American Community Survey 2019. As shown, both rich and poor neighborhoods can be multi-modal deserts, and more than half of multi-modal deserts are located in areas with better socioeconomic status. Multi-modal deserts located near central Downtown are relatively poor, and multi-modal deserts located near the edge of Downtown are neighborhoods with relatively higher socioeconomic status. A detailed analysis of these multi-modal deserts is presented in the next section.

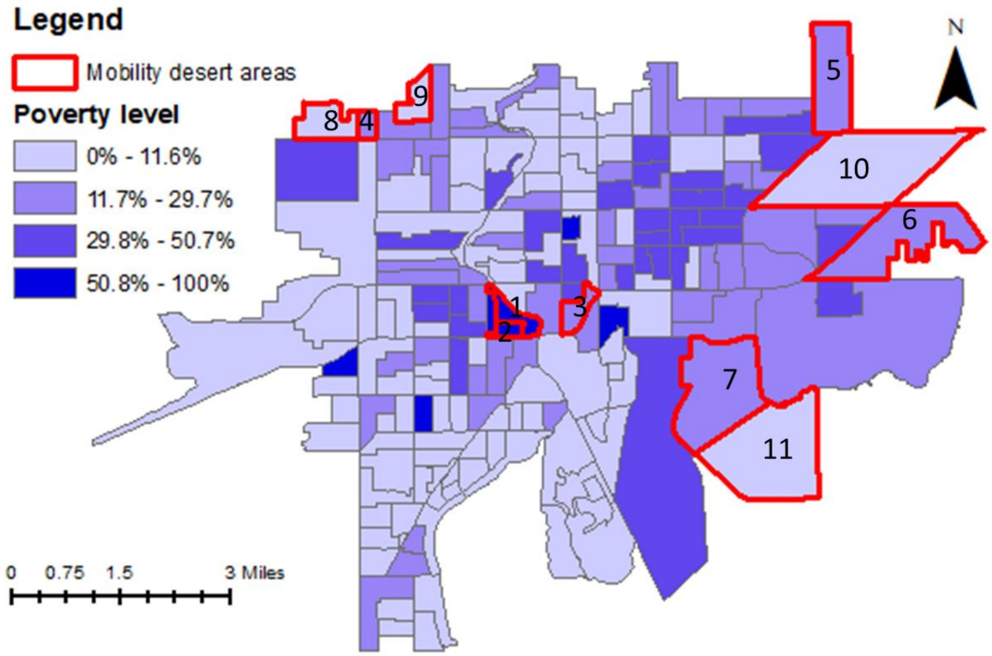


FIGURE 4: Multi-modal deserts in Downtown Tampa

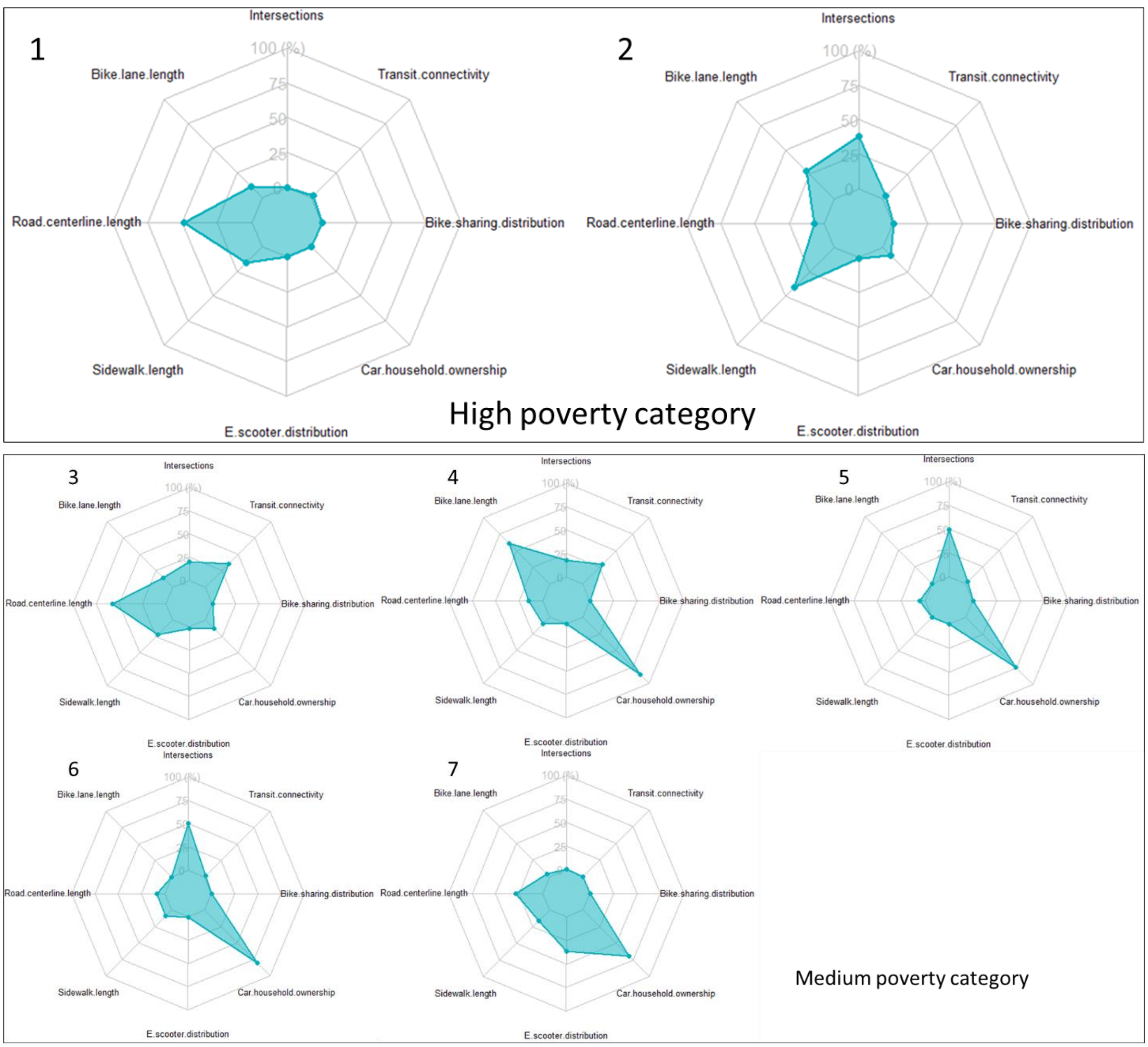
Discussion and policy implications

To inform practice and to better understand features that are correlated with multi-modal deserts, the features of each multi-modal desert were plotted in spider charts and are presented in Figure 5, corresponding to the 11 multi-modal deserts identified in Figure 4. In the spider chart, the value of each feature ranges from 0% to 100%, marked at intervals of 25%, 50%, and 75%. Based on socioeconomic status, multi-modal deserts were categorized into three groups, noted as High (greater than 50.8%), Medium (between 11.7% to 50.7%), and Low poverty ratios (lower than 11.6%).

For multi-modal deserts located in areas with a high poverty ratio, two areas were identified and are located in neighborhoods near central Downtown, shown in Figure 5 as Desert 1 and Desert 2. Neighborhoods located in the outskirts of central Downtown are of a historically high poverty ratio, and gentrification of the city has shifted most of the disadvantaged areas north toward the University area (*53*). Despite local authorities having spent much effort to improve mobility services, challenges remain—households have a low level of car ownership due to poverty, and the spatial coverage of shared micromobility options fail to reach these areas. Road infrastructure such as sidewalks and bike lanes are generally of low quality, and many streets are not well-connected; thus, households in these two Block Groups face challenges in receiving mobility services. These areas need special attention from policymakers, as the households are generally poor and the mobility constraints suffered in these neighborhoods further discourage residents from accessing critical activities, jobs, and educational opportunities to improve their quality of life.

For multi-modal deserts located in areas with a medium poverty ratio, one is in central Downtown, and others are located at the edge of Downtown. Desert 3 is the only one located in central Downtown, similar to zones with a high poverty ratio, Desert 3 has low mobility, which is mainly constrained by low vehicle ownership and limited infrastructure. It has some transit service, as it is located in central Downtown, although no shared micromobility is distributed, as they mainly operate in central Downtown areas. Other deserts are located at the edge of Downtown; these areas have relatively higher vehicle ownership and present some infrastructure such as bike lanes (Desert 4), road intersections (Desert 5 and 6), and e-scooters (Desert 7). Though Desert 7 has shared e-scooter services, the limited bike lane length cannot support this mode very well. Overall, these areas lack non-motorized infrastructure and transit services. Lack of access to public transit prevents low-income households from seeking job opportunities (*54*).

Multi-modal deserts with a low poverty ratio are located at the edge of Downtown. Given the higher socioeconomic status of these zones, all have a high vehicle ownership but a limited supply of infrastructure for non-motorized transportation and limited access to transit. People in these areas might not be constrained by auto travel; however, the lack of public transit encourages them to be more auto-dependent and discourages sustainable travel. If bike lanes and sidewalks were better developed, they could encourage more physical activity by using active transportation (*23*).



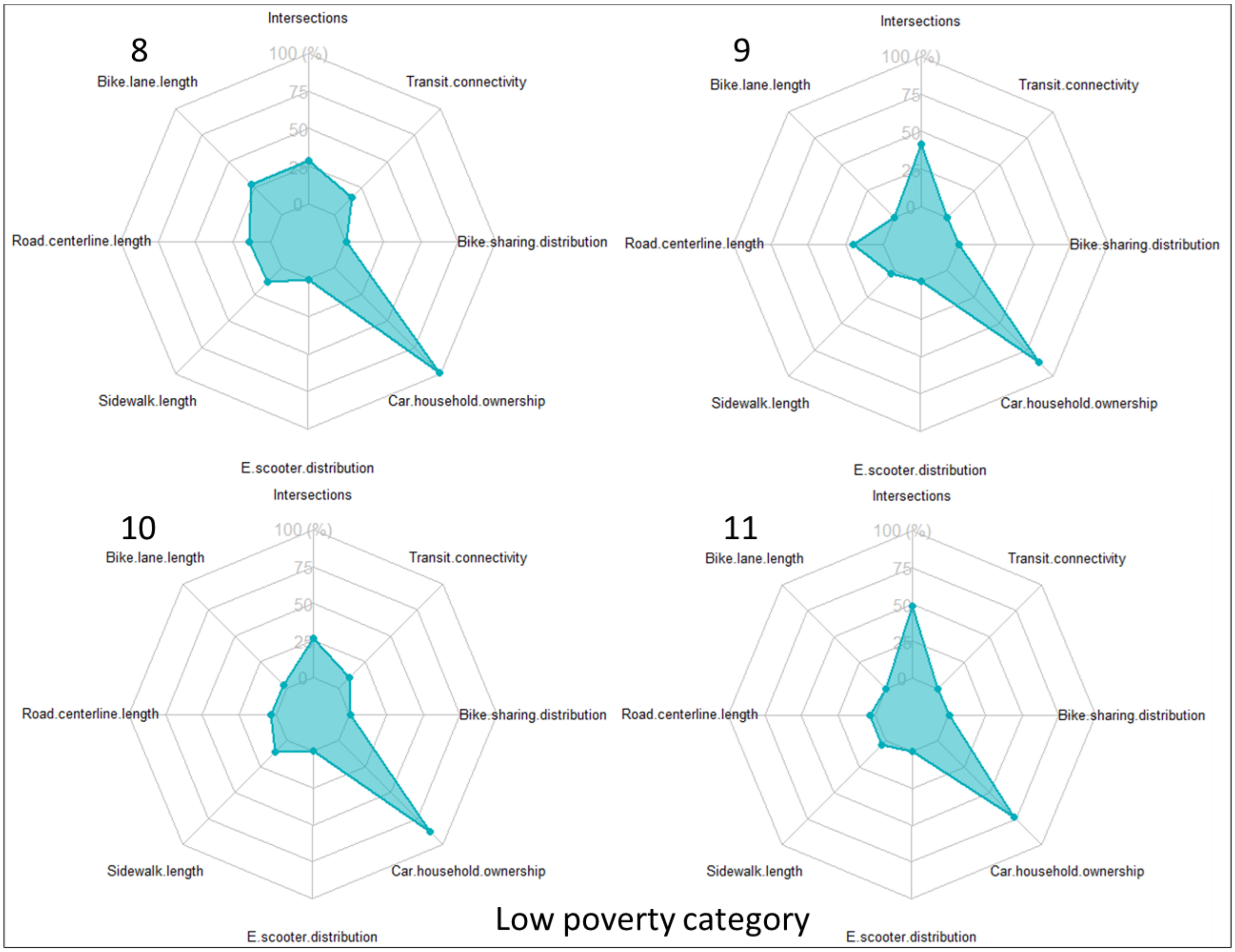


FIGURE 5: Spider charts of multi-modal deserts

To summarize, the identified multi-modal deserts show different patterns in Downtown Tampa. Multi-modal deserts with low-socioeconomic status are located near the Downtown core and are more constrained for all travel choices. Multi-modal deserts at the edge of Downtown generally have a high level of car dependency, and these areas are labeled as multi-modal deserts due to lack of infrastructure for active transportation, unconnected streets, and difficulty in accessing transit services and shared micromobility. Compared to the existing literature, results from this study provide additional insights towards transportation equity and planning. Previous studies have shown that shared micromobility targets specific populations with higher incomes and more education (*4; 8*) and are popular in areas with denser population (*55*). This may not be intentional, as most populated areas generally attract more users, incur more trips, and have a higher frequency of using public bikes and e-scooters (*56*). Such practices are related to identified multi-modal deserts in the outskirt areas. Figure 6 shows multi-modal deserts without considering shared micromobility; the input for multivariate outlier detection excludes the distribution of shared e-scooters and shared bikes. In this scenario, multi-modal deserts are mostly concentrated near the Downtown core. This indicates that areas that have no access to shared micromobility options are relatively disadvantaged in the previous analysis—the results can be greatly changed with different consideration of travel modes. An increasing spatial coverage and promoting equitable distribution of shared micromobility options can help mitigate challenges resulting from limited mobility services.

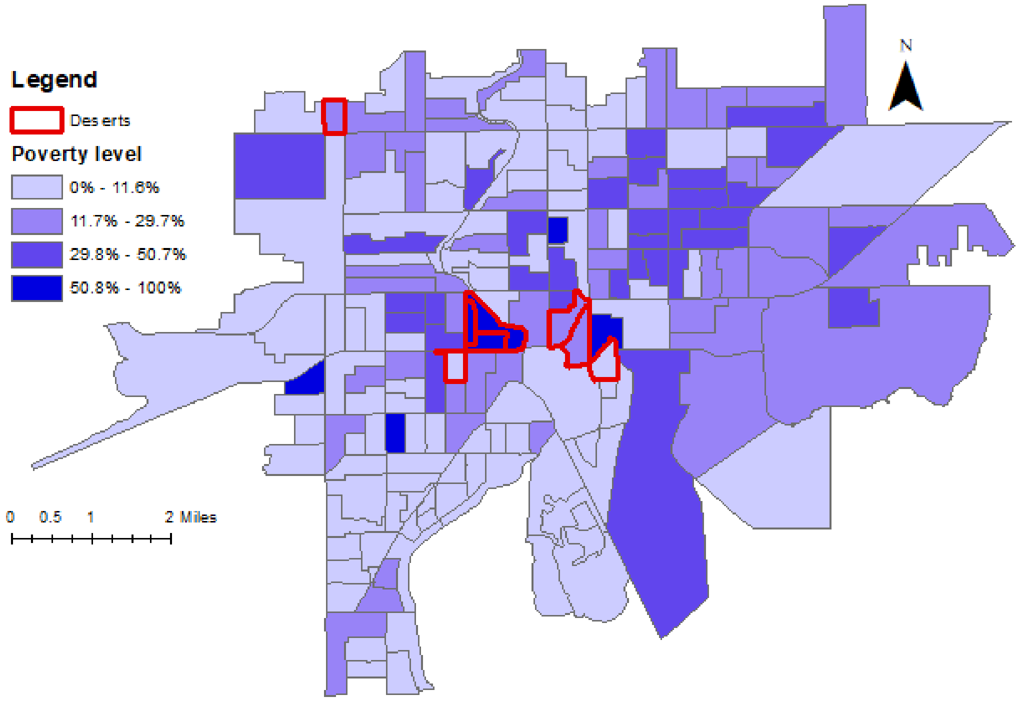


FIGURE 6: Multi-modal deserts without considering shared micromobility

To address multi-modal deserts in Downtown Tampa, several practices could be implemented:

- **Infrastructure** – The identified multi-modal deserts generally lack non-motorized infrastructure, which greatly reduces the feasibility of implementing related plans and discourages individuals from using sustainable alternatives. Policymakers should continually provide non-motorized infrastructure to promote multimodal transportation and ensure street connectivity, spatial coverage, and safety.
- **Shared micromobility** – Shared micromobility is an efficient way of providing alternatives. Unlike planning for bus routes, which is highly dependent on density and requires more investment, shared micromobility options can be deployed to disadvantaged areas quickly. In practice, policymakers could encourage shared micromobility companies to better serve disadvantaged neighborhoods (Desert 1 and 2) by reducing e-scooter application/permit fees; this practice has been implemented in cities such as Portland (*9*); findings suggest a positive effect on overall e-scooter use. To ensure affordability by low-income households, incentives and subsidies should be optimally provided to low-income users.
- **Transit** – Increasing the spatial coverage and frequency of transit services depends on many external factors, but a well-functioning transit network is important for reducing multi-modal deserts. Policymakers should investigate the demand for transit, specifically paying attention to disadvantaged groups; however, encouraging more compact development is key to a successful transit system.

CONCLUSIONS

Mobility is essential for people to participate in socioeconomic activities and obtain opportunities. The inequitable distribution of mobility resources leaves some areas lacking services, resulting in negative impacts on disadvantaged populations. To advocate for transportation equity, this study proposed the term “multi-modal desert” to describe areas with limited access to mobility services. A methodological framework was developed to identify multi-modal deserts and adopts an outlier detection method that uses robust PCA and distance measures to identify them. This can be used by policymakers to identify disadvantaged communities and develop solutions to optimally satisfy their mobility needs.

By applying the described methodological framework, this research successfully identified several multi-modal deserts. One group of multi-modal deserts has a high poverty ratio and poor infrastructure and does not provide easy access to any travel modes. These areas need special attention from policymakers; a systematic transportation planning approach could address the needs. Other multi-modal deserts have relatively higher degrees of vehicle ownership but generally lack the infrastructure to support active transportation. To promote health and sustainability, more bike lanes and sidewalks are needed to promote non-motorized travel in these areas. Transit services are generally poor in all identified multi-modal deserts. However, successful transit operation is conditioned on density. Future planning could initiate more high-density developments into such areas to keep a balance between promoting equity and ensuring the cost-effectiveness of the use of public funds.

This research could be extended in several aspects. First, this study adopted the Mahalanobis distance-based outlier detection; other methods such as k-nearest neighbor (*57*) and k-means (*58*) could be used, and results obtained from different analyses could be compared to enhance the robustness of the results. Second, future study could include more travel modes such as TNCs and paratransit, which are also popular options in metropolitan areas. This research did not include them due to lack of data, which may produce biased results; future research can incorporate TNCs and paratransit if the data are available and accessible. Third, this study considers transportation supply level/quality as a measure for transportation accessibility; other definitions of transportation accessibility that consider travel activities or affordability of using a transportation mode could be considered, especially when the result is targeted for interpretation for equity implications (*59*). Fourth, this study uses walk length and bike lane length to indicate the level of active transportation; there have been composite indices (*60-62*) for measuring how walkable/ bikeable a community is, and these indices are more comprehensive but require additional data sources (e.g., network distance to diverse destinations) for computation. Future studies could experiment with these indices if additional data can be obtained.

ACKNOWLEDGMENT

The authors gratefully acknowledge support provided by the Center for Transportation, Environment, and Community Health (CTECH), University Transportation Centers sponsored by the US Department of Transportation through Grant No. 69A3551747119. The contents of this manuscript reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. The contents do not necessarily reflect the official views or policies of US Department of Transportation.

REFERENCES

[1] Wellman, G. C. Transportation apartheid: the role of transportation policy in societal inequality. *Public Works Management & Policy,* Vol. 19, No. 4, 2014, pp. 334-339.
[2] NACTO. *Shared micromobility in the U.S.:2019*. https://nacto.org/shared-micromobility-2019/. Accessed 10/20, 2020.
[3] Jiang, J. *More Americans are using ride hailing apps*. https://www.pewresearch.org/fact-tank/2019/01/04/more-americans-are-using-ride-hailing-apps/. Accessed 11/8, 2019.
[4] Ursaki, J., and L. Aultman-Hall. Quantifying the equity of bikeshare access in US cities.In, University of Vermont. Transportation Research Center, 2015.
[5] Brown, A. E. Ridehail revolution: Ridehail travel and equity in Los Angeles.In, UCLA, 2018.
[6] Guo, Y., Z. Chen, A. Stuart, X. Li, and Y. Zhang. A systematic overview of transportation equity in terms of accessibility, traffic emissions, and safety outcomes: From conventional to emerging technologies. *Transportation research interdisciplinary perspectives,* Vol. 4, 2020, p. 100091.
[7] Welch, T. F., and S. Mishra. A measure of equity for public transit connectivity. *Journal of Transport Geography,* Vol. 33, 2013, pp. 29-41.
[8] Gavin, K., A. Bennett, A. H. Auchincloss, and A. Katenta. A brief study exploring social equity within bicycle share programs. *Transportation Letters,* Vol. 8, No. 3, 2016, pp. 177-180.
[9] PBOT. 2018 E-Scooter Findings Report.In, 2018.
[10] Young, M., C. Penny, and T. Ngongang. Baltimore City Dockless Vehicle Pilot Porgram Evaluation Report.In, Baltimore City DOT, 2019.
[11] Fortunati, J. *Mobility doesn't mean the same thing as transportation*. https://mobilitylab.org/2018/07/26/what-is-mobility/. Accessed 6/9, 2021.
[12] Van Wee, B., M. Hagoort, and J. A. Annema. Accessibility measures with competition. *Journal of Transport geography,* Vol. 9, No. 3, 2001, pp. 199-208.
[13] Jiao, J., and M. Dillivan. Transit deserts: The gap between demand and supply. *Journal of Public Transportation,* Vol. 16, No. 3, 2013, p. 2.
[14] Cai, M., J. Jiao, M. Luo, and Y. Liu. Identifying transit deserts for low-income commuters in Wuhan Metropolitan Area, China. *Transportation Research Part D: Transport and Environment,* Vol. 82, 2020, p. 102292.
[15] Jiao, J., and M. Cai. Using Open Source Data to Identify Transit Deserts in Four Major Chinese Cities. *ISPRS International Journal of Geo-Information,* Vol. 9, No. 2, 2020, p. 100.
[16] Di Ciommo, F., and Y. Shiftan. Transport equity analysis.In, Taylor & Francis, 2017.
[17] Sawicki, D. S., and M. Moody. Developing transportation alternatives for welfare recipients moving to work. *Journal of the American Planning Association,* Vol. 66, No. 3, 2000, pp. 306-318.
[18] Smart, M. J., N. J. Klein, and M. N. T. R. Consortium. A longitudinal analysis of cars, transit, and employment outcomes.In, Mineta National Transit Research Consortium, 2015.
[19] Mishra, S., T. F. Welch, and M. K. Jha. Performance indicators for public transit connectivity in multi-modal transportation networks. *Transportation Research Part A: Policy and Practice,* Vol. 46, No. 7, 2012, pp. 1066-1085.
[20] Couch, S., and H. K. Smalley. Encouraging Equitable Bikeshare: Implications of Docked and Dockless Models for Spatial Equity. *arXiv preprint arXiv:1906.00129*, 2019.
[21] Mooney, S. J., K. Hosford, B. Howe, A. Yan, M. Winters, A. Bassok, and J. A. Hirsch. Freedom from the station: Spatial equity in access to dockless bike share. *Journal of Transport Geography,* Vol. 74, 2019, pp. 91-96 %@ 0966-6923.

[22] Shellooe, S. D. Wheels When Who Wants Them: Assessing Social Equity and Access Implications of Carsharing in NYC.In, Columbia University, 2013.
[23] Saelens, B. E., J. F. Sallis, and L. D. Frank. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. *Annals of behavioral medicine,* Vol. 25, No. 2, 2003, pp. 80-91.
[24] Bao, J., X. Shi, and H. Zhang. Spatial Analysis of Bikeshare Ridership With Smart Card and POI Data Using Geographically Weighted Regression Method. *IEEE Access,* Vol. 6, 2018, pp. 76049-76059 %@ 72169-73536.
[25] Dill, J. Bicycling for transportation and health: the role of infrastructure. *Journal of public health policy,* Vol. 30, No. 1, 2009, pp. S95-S110.
[26] Estupiñán, N., and D. A. Rodríguez. The relationship between urban form and station boardings for Bogota's BRT. *Transportation Research Part A: Policy and Practice,* Vol. 42, No. 2, 2008, pp. 296-306.
[27] Sreelekha, M., K. Krishnamurthy, and M. Anjaneyulu. Interaction between road network connectivity and spatial pattern. *Procedia technology,* Vol. 24, 2016, pp. 131-139.
[28] Jarrell, M. G. A comparison of two procedures, the Mahalanobis distance and the Andrews-Pregibon statistic, for identifying multivariate outliers. 1992.
[29] Rasmussen, J. L. Evaluating outlier identification tests: Mahalanobis D squared and Comrey Dk. *Multivariate behavioral research,* Vol. 23, No. 2, 1988, pp. 189-202.
[30] Stevens, J. P. Outliers and influential data points in regression analysis. *Psychological bulletin,* Vol. 95, No. 2, 1984, p. 334.
[31] Anscombe, F. J. Rejection of outliers. *Technometrics,* Vol. 2, No. 2, 1960, pp. 123-146.
[32] Osborne, J. W., and A. Overbay. The power of outliers (and why researchers should always check for them). *Practical Assessment, Research, and Evaluation,* Vol. 9, No. 1, 2004, p. 6.
[33] Tukey, J. W. *Exploratory data analysis*. Reading, Mass., 1977.
[34] Van der Loo, M. P. *Distribution based outlier detection in univariate data*. Statistics Netherlands, 2010.
[35] Hodge, V., and J. Austin. A survey of outlier detection methodologies. *Artificial intelligence review,* Vol. 22, No. 2, 2004, pp. 85-126.
[36] Smiti, A. A critical overview of outlier detection methods. *Computer Science Review,* Vol. 38, 2020, p. 100306.
[37] Wang, H., M. J. Bah, and M. Hammad. Progress in outlier detection techniques: A survey. *Ieee Access,* Vol. 7, 2019, pp. 107964-108000.
[38] Finch, H. Distribution of variables by method of outlier detection. *Frontiers in Psychology,* Vol. 3, 2012, p. 211.
[39] Sunderland, K. M., D. Beaton, J. Fraser, D. Kwan, P. M. McLaughlin, M. Montero-Odasso, A. J. Peltsch, F. Pieruccini-Faria, D. J. Sahlas, and R. H. Swartz. The utility of multivariate outlier detection techniques for data quality evaluation in large studies: an application within the ONDRI project. *BMC medical research methodology,* Vol. 19, No. 1, 2019, pp. 1-16.
[40] Mahalanobis, P. C. On the generalized distance in statistics.In, National Institute of Science of India, 1936.
[41] Brereton, R. G. The chi squared and multinormal distributions. *Journal of Chemometrics,* Vol. 9, 2014, pp. 9 - 12.
[42] Filzmoser, P., A. Ruiz-Gazen, and C. Thomas-Agnan. Identification of local multivariate outliers. *Statistical Papers,* Vol. 55, No. 1, 2014, pp. 29-47.

[43] Rousseeuw, P. J., and K. V. Driessen. A fast algorithm for the minimum covariance determinant estimator. *Technometrics,* Vol. 41, No. 3, 1999, pp. 212-223.
[44] Rousseeuw, P. J. Least median of squares regression. *Journal of the American statistical association,* Vol. 79, No. 388, 1984, pp. 871-880.
[45] ---. Multivariate estimation with high breakdown point. *Mathematical statistics and applications,* Vol. 8, No. 283-297, 1985, p. 37.
[46] Cohen Freue, G. V., Z. Hollander, E. Shen, R. H. Zamar, R. Balshaw, A. Scherer, B. McManus, P. Keown, W. R. McMaster, and R. T. Ng. MDQC: a new quality assessment method for microarrays based on quality control reports. *Bioinformatics,* Vol. 23, No. 23, 2007, pp. 3162-3169.
[47] Steinbuss, G., and K. Böhm. Hiding outliers in high-dimensional data spaces. *International Journal of Data Science and Analytics,* Vol. 4, No. 3, 2017, pp. 173-189.
[48] Jolliffe, I. T., and J. Cadima. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* Vol. 374, No. 2065, 2016, p. 20150202.
[49] Pascoal, C., M. R. Oliveira, A. Pacheco, and R. Valadas. Detection of outliers using robust principal component analysis: A simulation study.In *Combining Soft Computing and Statistical Methods in Data Analysis*, Springer, 2010. pp. 499-507.
[50] Hubert, M., P. J. Rousseeuw, and K. Vanden Branden. ROBPCA: a new approach to robust principal component analysis. *Technometrics,* Vol. 47, No. 1, 2005, pp. 64-79.
[51] Stahel, W. Robust estimation: Infinitesimal optimality and covariance matrix estimators, Zurich.In, Ph. D Dissertation, ETH, 1981.
[52] Donoho, D. L. Breakdown Properties of Multivariate Location Estimators.In, Ph.D. qualifying paper, Harvard University, 1982.
[53] Eagan, T. Concentrations of Poverty Hillsborough County: 1970 - 2018.In, Plan Hillsborough, 2020.
[54] Serulle, N. U., and C. Cirillo. Transportation needs of low income population: a policy analysis for the Washington DC metropolitan region. *Public Transport,* Vol. 8, No. 1, 2016, pp. 103-123.
[55] Krykewycz, G. R., C. M. Puchalsky, J. Rocks, B. Bonnette, and F. Jaskiewicz. Defining a primary market and estimating demand for major bicycle-sharing program in Philadelphia, Pennsylvania. *Transportation research record,* Vol. 2143, No. 1, 2010, pp. 117-124.
[56] Jiao, J., and S. Bai. Understanding the Shared E-scooter Travels in Austin, TX. *ISPRS International Journal of Geo-Information,* Vol. 9, No. 2, 2020, p. 135.
[57] Wettschereck, D. A study of distance-based machine learning algorithms. 1994.
[58] Chawla, S., and A. Gionis. k-means–: A unified approach to clustering and outlier detection.In *Proceedings of the 2013 SIAM International Conference on Data Mining*, SIAM, 2013. pp. 189-197.
[59] Litman, T. *Transportation affordability*. Victoria Transport Policy Institute, 2017.
[60] Frank, L. D., J. F. Sallis, B. E. Saelens, L. Leary, K. Cain, T. L. Conway, and P. M. Hess. The development of a walkability index: application to the Neighborhood Quality of Life Study. *British journal of sports medicine,* Vol. 44, No. 13, 2010, pp. 924-933.
[61] Porter, A. K., H. W. Kohl III, A. Pérez, B. Reininger, K. Pettee Gabriel, and D. Salvo. Bikeability: assessing the objectively measured environment in relation to recreation and transportation bicycling. *Environment and Behavior,* Vol. 52, No. 8, 2020, pp. 861-894.

[62] Carr, L. J., S. I. Dunsiger, and B. H. Marcus. Walk score™ as a global estimate of neighborhood walkability. *American journal of preventive medicine,* Vol. 39, No. 5, 2010, pp. 460-463.